Training native speakers of Japanese to distinguish English /ɹ/ and /l/: Order effect of the perception training and production training

Chika Fujiyuki^{1,*}, Misato Kitamura², Sayoko Eguchi^{1,2}, Shuyi Yang¹ and Reiko Akahane-Yamada^{1,2,†}

¹Graduate School of Intercultural Studies, Kobe University,
1-2-1, Tsurukabuto, Nada-ku, Kobe, 657-8501 Japan

²Advanced Telecommunication Research Laboratories (ATR),
2-2, Hikaridai, Seika, Soraku, Kyoto, 619-0288 Japan

(Received 16 January 2017, Accepted for publication 2 March 2017)

Keywords: Second language, Speech perception, Speech production, Laboratory training
PACS number: 43.71.Ft [doi:10.1250/ast.38.229]

1. Introduction

It is not always easy for adult learners to enhance their listening and speaking skills of a second language (L2), partly because their perception and production of L2 are strongly affected by the phonological system of their first language (L1). Previous research has shown that foreign accent persists even for highly proficient speakers of a non-native language [e.g. 1, 2]. Laboratory training studies in the past decades have demonstrated that adult learners can overcome the negative influence of L1 phonological system and enhance L2 phonological skills through intensive training or training on a daily basis.

A series of laboratory training studies which trained native speakers of Japanese to perceptually identify English /ɹ/ and /l/ showed that an extensive training using high variability approach, i.e. using word stimuli contrasted in /ɹ/ and /l/ in multiple phonetic contexts and produced by multiple talkers, improved trainees' identification accuracy from pretest to posttest, and the training effect generalized to untrained words and untrained talker [e.g. 3]. Furthermore, it was found that the perception learning will bring improvement in production [4]. Wang *et al.* also showed that American speakers improved in producing four tones of Mandarin Chinese after being trained to perceptually identify these contrasts [5]. Akahane-Yamada *et al.* trained Japanese speakers to produce English /ɹ/ and /l/ by giving spectrographic representation or HMM-based pronunciation scores as feedback and found that trainees have improved both their production and perception [6,7]. These studies indicate that the effect of perception training transfers to production domain, and vice versa.

Baese-Berk and Samuel reported on an intriguing result of a perception-production study [8]. They compared perception-only training and perception plus production training and found that producing tokens during training can disrupt perceptual learning.

We now further investigate into the relationship between perception and production in the scope of developing an effective training program. In this study, we attempted to

optimize perception and production training of English /ɹ/-/l/ contrast by examining order effect. This paper will report on the result of the perception test and a conclusion we drew from the result.

2. Method

2.1. Participants

Thirty-four native speakers of Japanese (19 males and 15 females), aged between 20 to 30 years with no speech or hearing impairment, participated the study. They all speak Japanese as their first language and have no experience of being in English speaking countries more than 6 months. They were divided into three groups based on their achievement in the pretest: 1. First group received perception training first, then production training; 2. second group received production training first, then perception training; 3. third group received perception and production trainings alternately (c.f. Table 1). Pre-, mid-, and posttest were administered to all the participants to examine training effect. Both tests and training were self-paced and conducted in a quiet room using laptops. Participants were instructed to take breaks whenever necessary.

2.2. Stimuli

Fifty English minimal pairs contrasted in /ɹ/ and /l/ produced by a native speaker of American English were used in the test and 49 pairs produced by another five native speakers of American English were used in the training. Twenty-four pairs appeared in both tests and training. The contrast appeared in five different phonetic environments; word-initial singleton, word-initial consonant cluster, intervocalic, word-final, and word-final consonant cluster. The presentation of stimuli was made through a headset.

2.3. Perception training

A two alternative forced choice (2AFC) task was used. In each trial, both members of a minimal pair in standard English orthographic form was shown on the laptop monitor, and either one of the pair was played over the headphone. Participants chose the word they heard by pressing a button. Feedback was given in form of a chime (correct) or a buzzer (incorrect) on each response. Each perception training session consisted of 980 trials (49 minimal pairs, i.e., 98 words, each spoken by five different speakers, presented twice within a

*e-mail: 141c306c@stu.kobe-u.ac.jp

†e-mail: yamada@atr-lt.jp

Table 1 Training schedule of each group.

Day	1 st group	2 nd group	3 rd group
1	Pretest	Pretest	Pretest
2	Perception training (2 sessions)	Production training (0.5 session)	Perception training (1.2 session) Production training (0.2 session)
3	Perception training (2 sessions)	Production training (0.5 session)	Perception training (1.2 session) Production training (0.2 session)
4	Perception training (2 sessions)	midtest	Perception training (0.6 session) Production training (0.1 session)
5	midtest	Perception training (2 sessions)	midtest Perception training (0.6 session) Production training (0.1 session)
6	Production training (0.5 session)	Perception training (2 sessions)	Perception training (1.2 session) Production training (0.2 session)
7	Production training (0.5 session)	Perception training (2 sessions)	Perception training (1.2 session) Production training (0.2 session)
8	Posttest	Posttest	Posttest

session), and lasted approximately 75 minutes. Each participant received two sessions (1960 trials, ca. 150 minutes) per day, six sessions (5,880 trials, ca. 450 minutes) in total.

2.4. Production training

The production training consisted of two different tasks: repetition task and translation - and - production task. In the repetition task, participants were asked to pronounce the word displayed on the screen following model pronunciation. In this task, participants' production in each trial was automatically analyzed using pronunciation scoring technique (from 0 to 100 points, with 100 being the best) and participants were allowed to move on to the next trial when achieving a score higher than 50 points or having repeated the same trail twice. In the translation - and - production task, participants were asked to choose one of the minimal pair that is an English equivalent of the Japanese word displayed on the screen, and pronounce the word. In this task, feedback was given in form of a chime (correct) or a buzzer (incorrect) on each response. One session of the production training consisted of 980 trials, lasted about 200 minutes. Participants received a half of a session a day, one session in total.

2.5. Test

Participants received the perception and the production tests before (pretest), during (midtest), and immediately after (posttest) the training period. A 2AFC task, the same task used in the perception training, was used in the perception test. No feedback was given. The perception test consisted of 50 minimal pairs. The repetition task and the translation-and-production task, identical to those in the production training, were used in the production test. No feedback was given in both tasks. In the repetition task, participants' production was

Fig. 1 Correct response rate in the perception test for each group at pre-, mid-, and posttest.

recorded through a headset and saved on a laptop. A task in each test consisted of 50 trials, i.e., a participant received 150 trials in total in each test phase.

3. Result of perception test

The results of the perception test for all three groups are displayed in Fig. 1. All groups improved their perception skill from pretest to posttest; an increase of 15.5 was found for the first group (perception→production), 10.2 for the second group (production→perception) and 12.8 for the third group (mix). A two-way ANOVA with group (first, second, third) and test phase (pre, mid, post) as variables showed a significant main effect of test phase [$F(2, 31) = 56.451$, $p < 0.001$] and a significant group and test phase interaction [$F(4, 62) = 4.167$, $p < 0.01$]. Post hoc pairwise comparisons with Ryan's method (Fig. 2) showed a significant improvement between pretest and posttest for all three groups ($p < 0.001$), between pretest and midtest for the first group ($p < 0.001$) and the third group ($p < 0.01$), from midtest to posttest for the second group ($p < 0.001$) and the third group ($p < 0.01$). A significant difference was observed at midtest, between the first and the second group ($p < 0.001$) as well as between the first and the third group ($p < 0.01$). There was no difference among three groups at pretest and posttest. The result and the analysis of the production test is still on progress and will be reported in our forthcoming report.

4. Discussion

The result of the perception test demonstrates significant training effect for all three groups. A remarkable improvement found between pre- and midtest for the first group should result from two facts: 1. this group received the perception training between the two tests; 2. this group completed more trials and spent more time in training compared to the other groups during the period. However, our concern in this study is to examine order effect. No significant difference among groups at both pretest and posttest suggests that order effect was not evident, at least in the perception ability. In the study of Yang and Akahane-Yamada [9], where Japanese speakers with no prior Chinese knowledge received perception and

Fig. 2 The result of post hoc comparison for the perception test (** $p < 0.01$).

vocabulary training on identifying four Mandarin tones, order effect was evident along with successful learning in both trainings: The group which started with the perception training before the vocabulary training achieved a greater improvement in some of the tests. They also found that order effect was more prominent in the perception tests than the vocabulary tests. Eguchi and Akahane-Yamada examined order effect in pronunciation and vocabulary training using English low-familiarity words [10]. Their study demonstrated successful learning in both trainings, no order effect, however, was observed. Taking all results together, we speculate that order effect is only evident in two trainings that deal with items from different linguistic levels (e.g. phonemes and lexicons). It was also indicated that experience with the target language plays a role in order effect: Order effect may be positive only in training at early stages of second language learning.

We are going to discuss further on the issue, including the result and analysis of the production test, in our forthcoming report.

References

- [1] S. Tahta, M. Wood and K. Loewenthal, "Foreign accents: Factors relating to transfer of accent from the first language to a second language," *Lang. Speech*, **24**, 265–272 (1981).
- [2] J. E. Fledge and J. Hillenbrand, "Limits on phonetic accuracy in foreign language speech production," in *Interlanguage Phonology: The Acquisition of a Second Language Sound System*, G. Loup and S. Weinberger, Eds. (Newbury House, Cambridge, 1987), pp. 176–201.
- [3] S. E. Lively and D. B. Pisoni, "Training Japanese listeners to identify English /ɪ/ and /I/ II: The role of phonetic environment and talker variability in learning new perceptual categories," *J. Acoust. Soc. Am.*, **94**, 1242–1255 (1993).
- [4] A. R. Bradlow, D. B. Pisoni, A. R. Yamada and Y. Tohkura, "Training Japanese listeners to identify English /ɪ/ and /I/: IV. Some effects of perceptual learning on speech production," *J. Acoust. Soc. Am.*, **101**, 2299–2310 (1997).
- [5] Y. Wang, A. Jongman and J. A. Sereno, "Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training," *J. Acoust. Soc. Am.*, **113**, 1033–1043 (2003).
- [6] R. Akahane-Yamada, T. Adachi and H. Kawahara, "Second language production training using spectrographic representations as feedback," *J. Acoust. Soc. Jpn. (E)*, **18**, 341–343 (1997).
- [7] R. Akahane-Yamada, T. Adachi, H. Kawahara, J. S. Pruitt and E. McDermott, "Toward the optimization of computer-based second language production training," *Proc. ESCA Workshop on Speech Technology in Language Learning (STiLL 98)*, pp. 111–114 (1998).
- [8] M. M. Baese-Berk and A. G. Samuel, "Listeners beware: Speech production may be bad for learning speech sounds," *J. Mem. Lang.*, **89**, 23–36 (2016).
- [9] S. Yang and R. Akahane-Yamada, "Effects of perception training and vocabulary training of Mandarin tones by native speakers of Japanese," to be published in *Chin. J. Phon.* (2017).
- [10] S. Eguchi and R. Akahane-Yamada, "nihongo bogo wo taishou to shita eitango gakusyu no sai no hatsuon kadai to taiyaku kadai no junjo kouka," *Proc. 30th General Meeting of the Phonetic Society of Japan*, pp. 26–31 (2016).